Theor Appl Genet (2007) 114:1029–1038
DOI 10.1007/s00122-006-0496-7

ORIGINAL PAPER

Microsatellite variation in *Avena sterilis* oat germplasm

Yong-Bi Fu · James Chong · Tom Fetch · Ming-Li Wang

Received: 28 August 2006 / Accepted: 21 December 2006 / Published online: 31 January 2007
© Springer-Verlag 2007

Abstract The *Avena sterilis* L. collection in the Plant Gene Resources of Canada (PGRC) consists of 11,235 accessions originating from 27 countries and is an invaluable source of genetic variation for genetic improvement of oats, but it has been inadequately characterized, particularly using molecular techniques. More than 35 accessions have been identified with genes for resistance to oat crown and stem rusts, but little is known about their comparative genetic diversity. This study attempted to characterize a structured sample of 369 accessions representing 26 countries and two specific groups with *Puccinia coronata avenae* (Pc) and *Puccinia graminis avenae* (Pg) resistance genes using microsatellite (SSR) markers. Screening of 230 SSR primer pairs developed from other major crop species yielded 26 informative primer pairs for this characterization. These 26 primer pairs were applied to screen all the samples and 125 detected alleles were scored for each accession. Analyses of the SSR data showed the effectiveness of the stratified sampling applied in capturing country-wise SSR variation. The frequencies of polymorphic alleles ranged from 0.01 to 0.99 and averaged 0.28. More than 90% of the SSR variation resided within accessions of a country. Accessions from Greece, Liberia, and Italy were genetically most diverse, while accessions from Egypt, Georgia, Ethiopia, Gibraltar, and Kenya were most distinct. Seven major clusters were identified, each consisting of accessions from multiple countries and specific groups, and these clusters were not well congruent with geographic origins. Accessions with Pc and Pg genes had similar levels of SSR variation, did not appear to cluster together, and were not associated with the other representative accessions. These SSR patterns are significant for understanding the progenitor species of cultivated oat, managing *A. sterilis* germplasm, and exploring new sources of genes for oat improvement.

Communicated by J.-L. Jannink.

Y.-B. Fu (✉)
Plant Gene Resources of Canada,
Saskatoon Research Centre,
Agriculture and Agri-Food Canada,
107 Science Place, Saskatoon S7N 0X2, SK, Canada
e-mail: fuy@agr.gc.ca

J. Chong · T. Fetch
Cereal Research Centre, Agriculture and Agri-Food Canada,
195 Dafoe Road, Winnipeg MB R3T 2M9, Canada

M.-L. Wang
USDA-PGRCU, 1109 Experiment Street,
Griffin 30223-1797, GA, USA

Introduction

Avena sterilis L., a wild progenitor of cultivated hexaploid oat (*A. sativa* L.), constitutes a large reservoir of genetic diversity for genetic improvement (Rajhathy and Thomas 1974; Coffman 1977; Frey 1991; Harder et al. 1992). This wild gene pool is known to possess several useful traits, including high growth rates (Takeda and Frey 1977), resistance to crown and stem rusts (Wahl 1958, 1970; Martens et al. 1980; Harder et al. 1990; Fox et al. 1997), resistance to nematodes (Clamot and Rivoal 1984), herbicide resistance (Somody et al. 1984), and high protein grain (Takeda and Frey 1977). Plant Gene Resources of Canada (PGRC; the Canadian national seed genebank) at Saskatoon currently

maintains the world's largest *A. sterilis* collection of 11,235 accessions originating from 27 countries (Diederichsen et al. 2001) and covering widely across the species range (Baum 1977). This collection, since established in 1960s, has contributed to the worldwide oat improvement with germplasm of genes for resistance to oat crown and stem rusts (Frey 1991; Harder et al. 1992; McMullen and Patterson 1992). More than 35 rust resistance genes have been identified from *A. sterilis* germplasm (Harder et al. 1992), and many of these genes have been introgressed into oat breeding programs (McKenzie et al. 1981, 1984; McKenzie and Harder 1995). However, this collection has not been adequately characterized, particularly using molecular techniques. No representative germplasm (or core collection) has been developed from the *A. sterilis* collection. Little is known about the comparative genetic diversity of the rust resistance accessions identified so far and their genetic associations with the entire collection.

Efforts have been made to characterize some germplasm of *A. sterilis*. Variation of six seed characters in 457 representative accessions of the USDA *A. sterilis* collection (largely originated from the PGRC collection; Goffreda et al. 1992) was enormous and not randomly distributed over the geographic area where they were collected (Rezai and Frey 1988). Isozyme analysis of 1,005 *A. sterilis* accessions from 23 countries established six broad genetic groups of *A. sterilis* (Phillips et al. 1993). Further isozyme analysis of accessions from Turkey also revealed a broad array of variation present in four major groups (Hunter et al. 1995). Restriction fragment length polymorphism (RFLP) analysis revealed a clear separation of *A. sterilis* accessions between the eastern (Iran and Iraq) and western (Algeria, Ethiopia, Israel, Lebanon, Morocco, and Syria) regions in a stratified sample of 173 *A. sterilis* accessions originating in Africa and Southwest Asia (Goffreda et al. 1992). These characterizations have enhanced our understanding of existing *A. sterilis* gene pool and facilitated germplasm selection for oat improvement (Murphy and Phillips 1993; Zhou et al. 1999). However, these findings offer limited assistance in managing the *A. sterilis* collection and in improving the search of useful new traits from *A. sterilis* germplasm.

Microsatellite (or simple sequence repeat; SSR) markers have proven to be important tools in crop genetics (Röder et al. 1998; Li et al. 2000; Gao et al. 2003; Yu et al. 2004) and germplasm research (Karp 2002; Fu et al. 2003). However, the number of SSR markers developed so far has been insufficient for cultivated oat research and much less for *A. sterilis* (Li et al. 2000; Holland et al. 2001; Pal et al. 2002; Jannink and Gardner 2005). Recent assessments on the transferability of SSR markers developed from wheat and barley to other grass species are encouraging (Wu et al. 2003; Saha et al. 2004; Wang et al. 2005), but no such assessments have been made in *A. sterilis* to enlarge the pool of SSR markers.

The objectives of this study were to (1) assess the genetic variation and structure of a stratified sample of *A. sterilis* accessions representing 26 countries and two specific groups of rust resistance accessions using SSR markers and (2) determine the genetic associations of rust resistance accessions with the other representative germplasm.

Materials and methods

Plant materials

A core subset of 339 accessions (Table 1) was developed from the PGRC *A. sterilis* collection through a random sampling stratified with respect to country of origin and with the size roughly equal to the natural logarithm frequency of accessions for a country (Brown 1989). Twenty-seven accessions with genes for resistance to *Puccinia coronata avenae* Cda. f. sp. *avenae* Eriks (Pc) and three accessions with genes for resistance to *Puccinia graminis* Pers. f. sp. *avenae* Eriks and E. Henn. (Pg) were selected from the accessions known with rust resistance to reflect the effort of identifying rust resistance genes over the last 60 years (Harder et al. 1992). The accessions with Pc resistance originated from eight countries, mostly from Israel and Algeria (Table 1), and contained the described resistance genes ranging from Pc35 to Pc98. The accessions with Pg resistance contained the genes for Pg13, Pg15, and Pg17, and these accessions originated from Spain, Tunisia, and Turkey. The detailed information on these rust resistance accessions is available upon request.

DNA extraction

About 10–15 kernels of each selected accession were obtained from the PGRC *A. sterilis* collection and grown in the greenhouse at the Saskatoon Research Centre, Agriculture and Agri-Food Canada. Young leaves were collected from ten 5-day-old seedlings of each accession, bulked, freeze-dried (in a Labconco Freeze Dry System for 3–5 days), and stored at −80°C. From each bulked sample, dry leaves were finely chopped and ground to a fine power in a 2-ml Eppen-

Table 1 Microsatellite variation for 369 *A. sterilis* accessions representing 26 countries

Country[a]	NA	NAS[b]	NPB	MBF	Range	WCV	CFst
Afghanistan	16	6	30	0.411	0.17–0.83	13.53	0.096
Algeria	405	20(9)	64	0.293	0.03–0.97	12.20	0.093
Cyprus	5	4	26	0.375	0.25–0.75	13.50	0.102
Egypt	3	3	8	0.458	0.33–0.67	5.33	0.140
Ethiopia	115	11	32	0.418	0.09–0.91	10.91	0.103
France	6	5	25	0.376	0.20–0.80	11.60	0.107
Georgia	5	5	22	0.334	0.13–0.87	8.47	0.117
Gibraltar	7	5	17	0.400	0.20–0.80	8.20	0.123
Greece	67	11	58	0.417	0.09–0.91	15.53	0.079
Iran	594	19	64	0.339	0.05–0.95	12.77	0.091
Iraq	96	22	67	0.349	0.05–0.95	13.39	0.087
Israel	6976	36(10)	79	0.350	0.02–0.98	13.09	0.088
Italy	79	14	53	0.298	0.07–0.93	14.09	0.085
Jordan	38	20	55	0.334	0.05–0.95	11.35	0.098
Kenya	17	10	36	0.397	0.10–0.90	13.67	0.090
Lebanon	63	12(1)	47	0.354	0.08–0.92	12.79	0.093
Liberia	56	11	51	0.362	0.09–0.91	14.98	0.082
Morocco	612	20(1)	46	0.206	0.05–0.95	10.79	0.101
Pakistan	3	3	13	0.385	0.33–0.67	9.33	0.124
Portugal	143	15(3)	54	0.374	0.06–0.94	12.91	0.090
Spain	458	22(1)(1)	60	0.308	0.04–0.96	12.25	0.093
Switzerland	36	9	31	0.337	0.11–0.89	10.06	0.109
Tunisia	364	21(1)(1)	50	0.271	0.04–0.96	10.26	0.104
Turkey	1034	21(1)(1)	65	0.364	0.04–0.96	12.35	0.093
UK	7	4	26	0.442	0.25–0.75	13.33	0.103
Yugoslavia*	29	10	38	0.287	0.10–0.90	10.36	0.107
Total	11234	369					
Average			43	0.355		11.81	0.100

NA Number of accessions held in the PGRC collection, *NAS* number of accessions assayed, *NPB* number of polymorphic bands, *MBF* mean band frequency, *Range* minimum and maximum band frequencies, *WCV* within country variation, *CFst* country specific *Fst* (proportion of variation)

[a] One accession from Ukraine was not included in this study due to insufficient seed amount. Yugoslavia* the former Yugoslavia

[b] The number of accessions selected for the core subset were shown without parenthesis, the number of accessions with Pc resistance genes in the first parenthesis, and the number of accessions with Pg resistance genes in the second parenthesis

dorf tube with two 3-mm glass beads on a horizontal shaker. Genomic DNA was extracted using DNeasy plant mini kit (Qiagen Inc., Mississauga, ON, Canada) according to the manufacturer's directions. Extracted DNA was quantified by fluorimetry using Hoechst 33258 stain (Sigma Chemical Co., St. Louis, MO, USA), followed by dilution to 25 ng/μl for SSR analysis.

SSR analysis

A set of 230 SSR primer pairs was selected from published genomic and expressed sequence tag (EST)-derived SSR primer pairs in oat (25 AM from Li et al. 2000), wheat (84 CWM from Gao et al. 2003, 65 KSUM and 22 CNL from Yu et al. 2004, and 22 GWM from Röder et al. 1998) and barley (12 HVM from Becker and Heun 1995 and Liu et al. 1996). Each primer pair was first screened on eight randomly selected accessions. All polymerase chain reactions (PCRs) were performed in an MJ Research DYAD thermocycler (BioRad, Mississauga, ON, Canada) using the original PCR conditions developed in oat (Li et al. 2000), wheat (Röder et al. 1998; Gao et al. 2003; Yu et al. 2004), and barley (Becker and Heun 1995; Liu et al. 1996). The PCR products were separated on a 6% (w/v) non-denaturing acrylamide:bis-acrylamide (19:1) gel in 1× TBE buffer with 0.5 mg/l ethidium bromide for 2–2.5 h and recorded on a digital gel documentation system. This screening revealed that 216 primer pairs produced some bands but 54 EST-derived and 28 genomic primer pairs showed only monomorphic bands. Further screening of the 82 (36 EST-derived and 46 genomic) polymorphic primer pairs with another eight samples confirmed the informativeness of 13 EST-derived and 13 genomic primer pairs in detecting polymorphism (Table 2). The selected primer pairs were applied to screen all 369 accessions. To minimize technique-born and scoring errors, three samples assayed in the first gel were randomly selected and duplicated in the other three gels.

Data analysis

To generate a dataset of SSR allele counts for each accession, DNA fragments amplified by SSR primer pairs were identified based on their sizes in base pairs measured with a 10 bp DNA ladder (Invitrogen, Carlsbad, CA, USA) and compared with the sizes reported in the literature. The scored alleles were first assessed for consistency with duplicated samples and alleles of a primer pair with 5% or higher error rates over the four gels were discarded. The levels of polymorphism were analyzed with respect to primer and country by counting the number of polymorphic bands and generating the summary statistics on the band frequencies. To assess the informativeness of each marker, the polymorphic information content (PIC) was calculated for each locus, as described in Roussel et al. (2004). To visualize the variation pattern, the numbers of polymorphic bands were plotted against their frequencies of occurrence in all the assayed accessions. To determine the effectiveness of sampling for each country, a regression was done using SAS PROC REG (SAS Institute 2004) on the number of accessions over the number of polymorphic bands, the mean band frequency, within-country variation, and country-specific proportion of variation obtained from the analysis of molecular variance (AMOVA; Excoffier et al. 1992) given below.

Table 2 Details of 26 microsatellite markers applied in this study

Primer pair	T/R[a]	Forward Primer Sequence (5′–3′)	Reverse Primer Sequence (5′–3′)	Repeat	Ta (°C)	NA[b]	ASR[c]	PIC[d]	Function[e]
CWM26	E/R1	GGCAGCAGCAAAAGCAGGTC	ACTCACAATAGGGCGAAATG	(A)46	55	4	35–44	0.04	s-MDH (Q08062)
CWM48	E/R1	GATCGGCGACTTCCTCCCTCAT	ACCCCGCTCTTTCCCCAATAAT	(ATG)7	55	2	135–300	0.87	WIVRV (AF069309)
CWM65	E/R1	TCATTGGTGTCATCCCTCGTGT	GAATAATGCCTTGACCCTGGAC	(AAC)6	55	2	45–48	0.02	Γ-gliadin (AAK84773)
CWM90	E/R1	GCAGCCATCCATAGCGTCGTG	AAATGGTGCCCTGATGATGGAG	(AAC)6	55	2	46–49	0.06	LMW-GS (BAB78749)
CWM162	E/R1	AACACAACATAGCGCAAGGAAGAT	GTTGCTGAGGCTGTTGGAAGGAGA	(AAC)11	55	1	80	0.09	α-gliadin (Q41530)
CWM163	E/R1	TTTCTCATCCTTGCCCTCGTG	CCTGGAAATTGTTGTTGTTGTA	(AAC)14	55	2	54–75	0.01	α –gliadin (AAB48476)
CWM204	E/R1	GCTACAAACCAGTCAGCAA	AACCGACCCTCCTCCTTC	(ATGT)6	55	1	204	0.85	
CWM214	E/R1	GCTCCTTGTTCACTCATCTC	ATGCAGTCCTACTTGGTGAT	(AG)18	50	2	48–54	ne(2)	LTI6B (BAC16385)
CWM340	E/R1	TGCTCCACGTGGCTGTAG	GTCGTCATCGTCATCGTCAT	(ACG)7	55	5	42–207	ne(4)	
HVMwG	E/R2	TCCAATGGCATCTACAGGACGGCCAA	GCAGGTTGAGCTGCGCAAAGTCGTCG	(AT)9	55	2	180–182	0.95	
KSUM37	E/R3	CCAACATTCAACATGCATCC	TTCGCGATGTCCAAACATAA	(CA)9	55	3	44–50	0.03	
KSUM69	E/R3	TCCCATCACCCATTTCTTCC	CCGTCGCTCATCTTCTTGCC	(CT)12	55	1	55	0.08	
KSUM176	E/R3	ATGCTCAAGCCGAGGAAGTA	ATGTCACCCATAGGAACGGA	(CTG)8	55	7	42–270	ne(4)	Protein (BE606965)
AM1	G/R4	GGATCCTCCACGCTGTTGA	CTCATCCGTATGGGCTTTA	(AG)21(CA GAG)6	58	11	150–240	0.41	
AM3	G/R4	CTGGTCATCCTCGCCGTTCA	CATTTAGCCAGGTTGCCAGGTC	(AG)35	58	24	144–378	0.93	
AM4	G/R4	GGTAAGGTTTCGAAGAGCAAAG	GGGCTATATCCATCCCTCAC	(AG)34	58	21	114–220	0.89	
AM5	G/R4	TTGTCAGCGAAATAAGCAGAGA	GAATTCGTGACCAGCAACAG	(AG)27	58	4	134–142	0.09	
AM19	G/R4	ATAGAACGGCATGATAACGAAATA	GCGCGACAACAGGACCTTC	(AC)3..(AC)6.. (AC)5..(AC)7	58	4	248–300	ne(3)	
AM22	G/R4	ATTGTATTTGTAGCCCCAGTTC	AAGAGCGACCCAGTTGTATG	(AC)22	58	3	60–146	0.34	
AM25	G/R4	AGCCTGGACATGTAATCTGGT	AGCCCTGGTCTTCTTCAACA	(AC)8.. (AC)4 (CT)4	58	3	220–240	0.10	
AM42	G/R4	GCTTCCCGCAAATCATCAT	GAGTAAGCAAAGGCCAAAAAGT	(GAA)16	58	6	188–201	0.40	
GWM3	G/R5	GCAGCGGCACTGGTACATTT	AATATCGCATCACTATCCCA	(CA)18	55	3	44–50	ne(3)	
GWM6	G/R5	CGTATCACCTCCTAGCTAAACTAG	AGCCTTATCATGACCCTACCTT	(GA)40	55	3	48–60	ne(3)	
GWM247	G/R5	GCAATCTTTTTTCTGACCACG	ATGTGCATGTCGGACGC	(GA)24	60	4	39–60	0.04	
GWM577	G/R5	ATGGCATAATTTGGTGAAATTG	TGTTTCAAGCCCAACTTCTATT	(CA)14(TA)6	55	2	50–55	ne(2)	
HVM3	G/R6	ACACCTTCCCAGGACAATCCATTG	AGCACGCAGAGCACCGAAAAAGTC	(AT)29	55	3	138–142	ne(3)	

[a] Type and reference; *E* EST-derived, *G* genomic, *R1* Gao et al. (2003), *R2* Becker and Heun (1995), *R3* Yu et al. (2004), *R4* Li et al. (2000), *R5* Röder et al. (1998), and *R6* Liu et al. (1996)

[b] *NA* Number of alleles observed

[c] *ASR* Allelic size range in base pair

[d] *PIC* Polymorphic information content, *ne* negative value and the number of possible alleles per sample given in the parenthesis

[e] Putative function and accession number shown in NCBI Genbank

Table 3 Analysis of molecular variance (AMOVA) sum of squares partitioning of total SSR variation into among- and within-country (or group) components based on several structure models

Structure model	*df*	Among-country (or group) component (%)[a]	Within-country (or group) component (%)
Core subset			
Country	25	10.03****	89.97
Core subset + specific groups			
Country	25	9.51****	90.49
Core subset versus specific groups			
Group	1	0.50*ns*	99.50

[a] Significance was tested with the probability that the among-country (or group) variance component was larger than zero, as computed from random permutations

**** and *ns* stand for the significance levels of $P < 0.0001$ and $P > 0.05$, respectively

To assess the genetic structure of 369 accessions representing various countries (or groups), an AMOVA was performed using Arlequin version 3.01 (Excoffier et al. 2005). This analysis not only allowed the partition of the total SSR variation into within- and among- country (or group) variation components, but also provided a measure of inter-country (or group) genetic distances as the proportion of the total SSR variation residing between oat accessions of any two countries (or groups) (called Phi statistic; Excoffier et al. 1992). Several structure models with country and specific group were applied (Table 3). Significance of resulting variance components and inter-country (or group) genetic distances was tested with 10,100 random permutations. To determine the genetic relationships of 369 accessions representing 26 countries, the inter-country distance matrices of Phi statistic were analyzed using NTSYS-pc 2.01 (Rohlf 1997) and clustered with the algorithm of unweighted pair-group methods using arithmetic averages.

To assess the genetic associations of individual accessions, a pairwise accession similarity matrix was generated using simple matching coefficient (Sokal and Michener 1958) and converted to the Euclidean distance matrix as the square root of 1 minus element-wise similarity for a principal coordinate analysis using NTSYS-pc program. The first three resulting principal coordinate scores were plotted to assess accession associations. A neighbor-joining analysis of 369 accessions was also made using PAUP* (Swofford 1998) and circulating tree was displayed using MEGA 3.01 (Kumar et al. 2004) to confirm the genetic association of individual accessions and to identify any genetic clustering without restriction to country of origin.

Results

The 26 SSR primer pairs applied in this study represented two types of SSR markers (i.e., 13 EST-derived and 13 genomic) developed in three crop species (16 wheat, 8 oat, and 2 barley). Specifically, they were consisted of five SSR types (1 single-, 15 di-, 8 tri-, 1 tetra-, and 1 penta-nucleotide repeats) and revealed at least 26 loci (Table 2). The number of alleles per primer pair ranged from 1 (CWM204) to 24 (AM3) and averaged 4.8 alleles. Allelic size varied greatly, ranging from 35 to 378 bp, depending on the length of repeat motif. PIC values of each primer pair ranged from 0.02 to 0.95 with an average of 0.35, and eight primer pairs had negative values with the number of possible alleles per sample ranging from two to four (Table 2). The most informative primer pair was HVMwG, followed by AM3, AM4, CWM48, and CWM204. The observed allelic frequencies ranged from 0.003 to 0.997 with an average of 0.276 (Fig. 1). Most of the alleles appear to be either frequently or infrequently present in the assayed accessions. Fifty-five alleles (44%) were present in a small proportion of the accessions (with frequencies of 0.05 or lower) and 27 alleles (22%) were present in most of the accessions (with frequencies of 0.86 or higher). Eight of the 13 EST-derived markers were associated with some putative functions, largely as proteins (Table 2). Thus, large SSR variations were observed in these *A. sterilis* accessions. (Fig. 2)

The core subset was selected by a stratified sampling of the *A. sterilis* collection with respect to country of origin. The number of accessions per country ranged from 3 (Egypt and Pakistan) to 33 (Israel) (Table 1). Such variable accession sizes were significantly correlated with the variation in the number of polymorphic bands, the mean band frequency, and the country-specific proportion of variation, but not with the within-country variation obtained from the AMOVA sum of squares (Table 1). Such patterns of correlation remain largely the same when 30 accessions of specific groups were analyzed in combination with the core subset. Genetic variation and structure as revealed from AMOVA were also similar for both core subset and

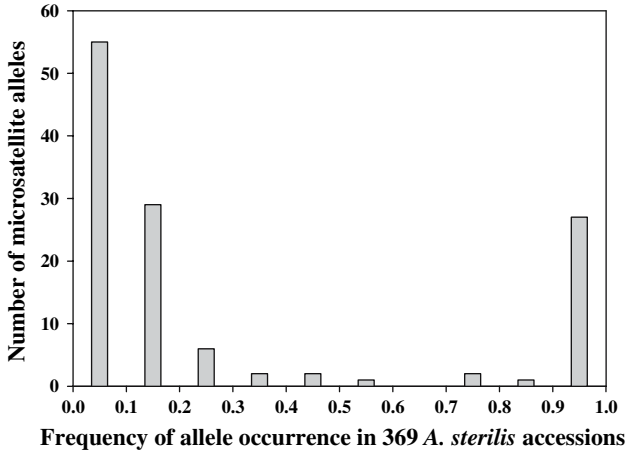

Fig. 1 Frequency distributions of 125 microsatellite alleles in 369 *A. sterilis* accessions

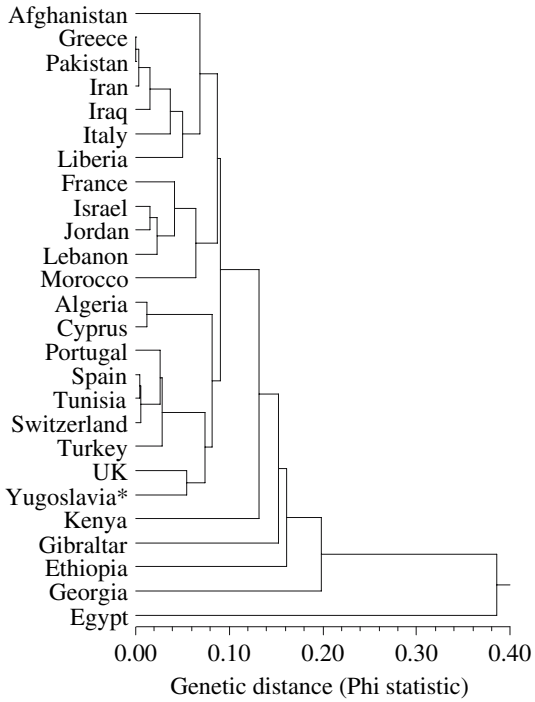

Fig. 2 Genetic associations of 369 *A. sterilis* accessions representing 26 countries. Yugoslavia* is the former Yugoslavia

specific groups (Table 3). For example, no difference was observed between the core subset and specific groups, although the proportion of variation explained by country was slightly reduced from 10.03 to 9.51% when the core subset was combined with the specific groups (Table 3). Because of these similarities, we will report the results from the combined dataset only.

Genetic variation for accessions of a country was estimated with four parameters. The number of polymorphic bands observed for accessions of a given country ranged from 8 to 79 with an average of 43. The mean band frequency ranged from 0.206 to 0.458 with an average of 0.355. The within-country variation ranged from 5.33 to 15.53 with an average of 11.81. The country-specific proportion of variation ranged from 0.079 to 0.140 with an average of 0.100. While these estimates are largely correlated, the estimate of within-country variation measures the amount of SSR variation residing in accessions of a country and the country-specific proportion of variation estimates the country-specific contribution to the overall SSR variation detected. The country with the most within-country variation was Greece (15.53), followed by Liberia (14.98) and Italy (14.09). The country with the least within-country variation was Egypt (5.33), followed by Gibraltar (8.20), Georgia (8.46), and Pakistan (9.33). These four countries contributed more to the overall SSR variation detected in this study as revealed by the country-specific proportions of variation (Table 1).

Assessment on the genetic relationships of the *A. sterilis* accessions representing 26 countries revealed five countries with the most distinct accessions and three major clusters. The countries with the most distinct accessions were Egypt, Georgia, Ethiopia, Gibraltar, and Kenya. The first major cluster consists of accessions from nine countries (Algeria, Cyprus, Portugal, Spain, Tunisia, Switzerland, Turkey, UK, and former Yugoslavia). The second cluster includes accessions from seven countries (Afghanistan, Greece, Pakistan, Iran, Iraq, Italy, and Liberia). The third group consists of accessions from five countries (France, Israel, Jordan, Lebanon, and Morocco). Such grouping does not seem to be correlated with geographic distribution of this species such as northern–southern or eastern–western separations, although accessions from Africa (such as Egypt, Ethiopia, and Kenya) appear to be genetically most distinct. Also, no marked associations of *A. sterilis* accessions with geographic regions can be visualized in Fig. 3, in which individual accessions of various origins were widely spread over the plot.

The neighbor-joining analysis revealed seven major clusters of 369 accessions as labeled in Fig. 4. These clusters vary in number of accessions, but consist of accessions from many countries or specific groups (Table 4). The largest cluster (IV) consists of 107 accessions originating from 23 countries, excluding Cyprus, France, and Switzerland (Table 4). The second largest cluster (VII) consists of 72 accessions from 15 countries. Cluster I consists of 12 genetically most distinct accessions originating from 10 countries (Table 4). Interestingly, these most distinct individual accessions did not originate from the countries with the most distinct accessions (i.e., Egypt, Georgia, Ethiopia, Gibraltar, and Kenya). Assessment on the accession

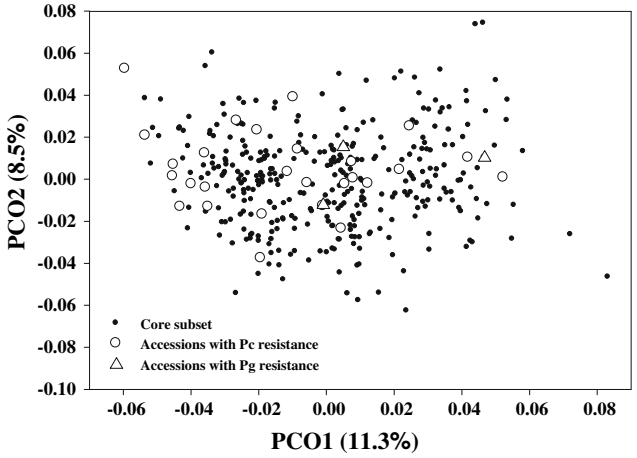

Fig. 3 Associations among 369 *A. sterilis* accessions, as revealed by a principal coordinate analysis. These two components accounted for 10.3 and 8.5% of the total variance, respectively. Germplasm of the core subset was labeled by *dot*, accessions with Pc resistance genes by *while circle*, and accessions with Pg resistance genes by *white triangle*

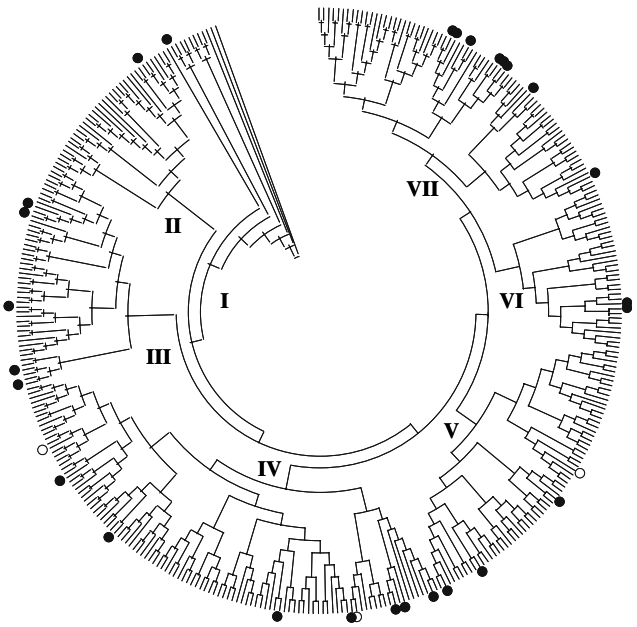

Fig. 4 Clustering of 369 *A. sterilis* accessions obtained from the neighbor-joining analysis of 125 microsatellite alleles. Seven major clusters were identified. Accessions with Pc and Pg genes were also marked with *black* and *white circles*, respectively

distributions in the seven clusters with respect to country of origin showed accessions of each country were spread in these clusters, depending on the number of accessions assayed for the country. Accessions from Israel, Spain, and Tunisia were represented in all seven clusters. Accessions from Portugal and Morocco resided in six of the seven clusters. Twenty-seven accessions with Pc genes were distributed in all seven clusters and three accessions with Pg genes were located in two clusters (IV and V; Fig. 4). No clusters were observed with respect to the period of resistance race identification. Such wide accession distributions of the Pc and Pg groups were also obvious in the principal coordinate plot (Fig. 3). No particular grouping, nor specific association with the core subset accessions, was observed for Pc and Pg groups.

Table 4 Accession distributions in seven major clusters obtained from the neighbor-joining analysis of 369 *A. sterilis* accessions with respect to country and specific group

Country or group	I	II	III	IV	V	VI	VII	Total
Afghanistan	1	1		2			2	6
Algeria			2	6	10	5	6	29
Cyprus			1		1		2	4
Egypt				3				3
Ethiopia			1	2	4	4		11
France	1		1			3		5
Georgia			1	2		2		5
Gibraltar		3		1			1	5
Greece			1	7	2	1		11
Iran			1	15		1	2	19
Iraq		1	3	15			3	22
Israel	1	6	12	5	11	2	9	46
Italy	1	2	1	7			3	14
Jordan			4	2	5	5	4	20
Kenya		6	2	1		1		10
Lebanon		1		4	3	2	3	13
Liberia	2		1	6	1		1	11
Morocco	1	3	5	6	2		4	21
Pakistan	1			2				3
Portugal		1	3	4	7	2	1	18
Spain	2	6	1	3	5	1	6	24
Switzerland					3	3	3	9
Tunisia	1	1	3	1	9	2	6	23
Turkey		2	4	8	4		5	23
UK	1			2			1	4
Yugoslavia*				3	5		2	10
Total	12	33	47	107	72	34	64	369
Pc group	1	1	5	7	3	2	8	27
Pg group				2	1			3

Yugoslavia* refers to the former Yugoslavia

Discussion

This study provides not only a new set of informative SSR markers for genetic studies of *A. sterilis*, but also a global view of *A. sterilis* genetic diversity. More than 90% of the SSR variation resided within accessions of a country. Accessions from Greece, Liberia, and Italy were genetically most diverse, while accessions from Egypt, Georgia, Ethiopia, Gibraltar, and Kenya were most distinct. Seven major clusters were identified, each consisting of accessions originating from multiple countries and specific groups, and these clusters were not well congruent with geographic origins. Accessions

with Pc and Pg resistance genes had similar levels of SSR variation, did not appear to cluster together, and were not associated with the other representative accessions. These SSR patterns are significant for understanding the progenitor species of cultivated oat, managing *A. sterilis* germplasm, and exploring new sources of genes for oat improvement.

While the transfer rate (57% = 132/230) of crop SSR markers to *A. sterilis* is compatible with those reported for other grass species (Wu et al. 2003; Saha et al. 2004; Wang et al. 2005), the effective transfer rate (11% = 26/230) is relatively low, largely because only scorable polymorphism was utilized for this study. The polymorphic bands scored in this study may have different fragment sizes from those previously reported. Also, some primer pairs revealed polymorphic bands probably at multiple loci in *A. sterilis*, rather than as reported from single loci in the crop species. This helps to explain, at least partly, the negative PIC values of the eight primer pairs. Thus, the chromosomal location of the markers reported here cannot be simply deduced from the linkage maps established in the crop species, and the exact coverage of the *A. sterilis* genome ($2n = 6\times = 42$) by these markers is unknown. These uncertainties make the genetic analysis of these markers for Hardy–Weinberg and linkage equilibriums less informative. Moreover, null alleles, if present, would be confounded with non-amplification in this study. These issues, although not surprising given the SSR transferability reported for other grass species (e.g. Wang et al. 2005), are worth further attention, especially for genetic mapping and association studies. Further assessment is warranted on the inheritance of the scored polymorphic bands and on the verification of SSR motifs by sequencing the scored bands. More effort is needed to enlarge the pool of SSR markers for oat research (Jannink and Gardner 2005).

This analysis also demonstrated the effectiveness of the stratified sampling in capturing the country-wise genetic variability from the whole collection, as evidenced in the significant associations of the variable accession sizes for different countries with the variations in the number of polymorphic bands, the mean band frequency, and the country-specific proportion of variation (Table 1). Thus, more country-wise variation was captured than would have been when using a completely random sampling scheme. This effectiveness was expected theoretically (Brown 1989), but it has not been well assessed empirically (Schoen and Brown 1993; Fu et al. 2005). However, it is not clear why the within-country variation was not correlated with the numbers of accessions from various countries. This is inconsistent with the results found in AFLP variation in cultivated hexaploid oat germplasm (Fu et al. 2005). The core subset assayed here represented only 3.3% of the *A. sterilis* accessions and theoretically could capture up to 70% of genetic diversity in the whole collection (Brown 1989). Thus, a considerable proportion of variation may still go undetected. Also, as a bulk sample for each accession was assayed, information on within-accession variation is not available and bias in detection of genetic variation could exist (Fu 2003). Moreover, effort was made to verify testing materials through passport data, but undetected contamination or misrepresentation (if any) could affect the comparisons of SSR variation. It is important to recognize these limitations in interpreting these results.

A large amount of genetic diversity was found in *A. sterilis* germplasm and more than 90% of the variation detected resided within the accessions of a country. These results are compatible with those variations based on seed characters (Rezai and Frey 1988), isozyme markers (Phillips et al. 1993; Hunter et al. 1995), and RFLP markers (Goffreda et al. 1992). Interestingly, the countries with the most within-country variations were Greece, Liberia, and Italy, but not those countries from the Middle East mentioned as oat diversity centers (Thomas 1995; Zeven and Zhukovsky 1975). The countries with the most distinct accessions were Egypt, Georgia, Ethiopia, Gibraltar, and Kenya. However, such distinctiveness could be questioned given the fact that the numbers of accessions available for these countries were small and the most distinct individual accessions did not originate from these five countries (Fig. 4; Table 4). Genetic clustering of *A. sterilis* accessions was not associated with geographic origins, and no clear separation of eastern–western accessions reported in the RFLP analysis (Goffreda et al. 1992) was found in this analysis. These findings, although less expected from a self-fertile species, are consistent with those obtained from isozyme analysis (Phillips et al. 1993) in which Turkey accessions were represented in all the six clusters identified. It is possible that the current geographic distribution of this species as displayed in our germplasm collection reflects only the dispersoion of *A. sterilis* established by human influence over thousand years, not its true weedy range (Baum 1977; Zohary and Hopf 2000). Comparison of these variation patterns with the AFLP variations reported in the core subset of cultivated hexaploid oat (Fu et al. 2005) revealed many similarities of variation pattern, which is well aligned with the view for *A. sterilis* being the progenitor of cultivated hexaploid oat (Thomas 1995; Zohary and Hopf 2000).

While the variation patterns obtained may be specific to this core subset, they should share some general

baseline information useful for the development of specific core subsets by determining the weighting or representation for each country (Brown and Spillane 1999). New core subsets should be developed with more emphasis given to accessions originating from countries with greater genetic distinctiveness such as Egypt, Ethiopia, and Kenya. Field collection or germplasm acquisition may also need to be directed toward these African countries to increase the coverage of the species distribution. Classifying accessions based on variation patterns of specific trait or disease data may be more informative than those based on geographical origin. Such stratification could enhance the genetic association studies between genetic markers and traits of interest based on linkage disequilibrium (Breseghello and Sorrells 2006).

Searching germplasm of particular traits would be greatly enhanced if genetic stratification existed in the germplasm collection and such stratification was associated with specific traits. However, the findings obtained in this study appear to indicate that genetic associations of individual accessions were not strong and accessions with Pc and Pg resistance genes were not associated with the representative accessions. This information helps to explain in part the difficulty in the previous identifications of *A. sterilis* germplasm with Pc and Pg genes. It also illustrates the challenges in future effort when searching germplasm of specific traits. For example, the genetic associations of rust resistance accessions reported here (Fig. 4; Table 4) do not support the previous expectation for searching for Pg genes from two relatively small geographic areas (the region around the western Black Sea and North Africa; Martens 1985). These findings may also suggest the exploitation of *A. sterilis* germplasm should be focused on germplasm within a country, as genetic stratification may exist within particular environmental niches. More fruitful genetic association studies of *A. sterilis* genes and SSR variation with respect to intra-country environment may need to be conducted within a specific genetic cluster and within a small geographic scale, particularly in countries like Turkey, Ethiopia, Israel and Morocco with tremendously contrasting environments within relatively short distances. Exploring new source of genes should be directed more toward accessions originating from Egypt, Ethiopia and Kenya with greater genetic distinctiveness.

Acknowledgments Authors would like to thank Ken Richards and Axel Diederichsen for their stimulating discussions on the project; Dallas Kessler, David Williams, Angela Taylor, and Gregory Peterson for their technical assistance in sampling, planting, and screening the germplasm; and Axel Diederichsen and David Williams for their helpful comments on the manuscript.

References

Baum BR (1977) Oats: wild and cultivated, a monograph of the Genus *Avena* L. (Poaceae). Thorn Press, Ottawa

Becker J, Heun M (1995) Barley microsatellites: allele variation and mapping. Plant Mol Biol 27:835–845

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177

Brown AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Brown AHD, Spillane C (1999) Implementing core collections—principles, procedures, progress, problems and promise. In: Johnson RC, Hodgkin T (eds) Collections for today and tomorrow. International Plant Genetic Resources Institute, Rome, pp 1–9

Clamot G, Rivoal R (1984) Genetic resistance to cereal cyst nematode *Heterodera avenae* Woll. in wild oat *Avena sterilis* I. 376. Euphytica 33:27–32

Coffman FA (1977) Oat history, identification and classification. USDA-ARS Tech Bull No 1516, Washington

Diederichsen A, Timmermans E, Williams DJ, Richards KW (2001) Holdings of *Avena* germplasm at Plant Gene Resources of Canada and status of the collection. Oat Newslett 47:35–42

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.01: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Fox SL, Brown PD, Chong J (1997) Inheritance of crown rust resistance in four accessions of *Avena sterilis* L. Crop Sci 37:342–345

Frey KJ (1991) Genetic resources of oats. In: Shands HL, Wiesner L (eds) Use of plant introduction in cultivar development. Crop Science Society of America Special Publication No. 17, Madison, pp 15–24

Fu YB (2003) Applications of bulking in molecular characterization of plant germplasm: a critical review. Plant Genet Res 1:161–167

Fu YB, Peterson GW, Scoles G, Rossnagel B, Schoen DJ, Richards KW (2003) Allelic diversity changes in 96 Canadian oat cultivars released from 1886 to 2001. Crop Sci 43:1989–1995

Fu YB, Peterson GW, Williams D, Richards KW, Mitchell Fetch J (2005) Patterns of AFLP variation in a core subset of cultivated hexaploid oat germplasm. Theor Appl Genet 111:530–539

Gao LF, Tang JF, Li HW, Jia JZ (2003) Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol Breed 12:245–261

Goffreda JC, Burnquist WB, Beer SC, Tanksley SD, Sorrells ME (1992) Application of molecular markers to assess genetic relationships among accessions of wild oat, *Avena sterilis*. Theor Appl Genet 85:146–151

Harder DE, Chong J, Brown PD, Martens JW (1990) Inheritance of resistance to *Puccinia coronata avenae* and *P. graminis avenae* in an accession of *Avena sterilis* from Spain. Genome 33:198–202

Harder DE, Chong J, Brown PD, Sebesta J, Fox S (1992) Wild oat as a source of disease resistance: history, utilization, and

prospects. In: Proceedings of the 4th international oat conference, wild oats in world agriculture, vol II. Adelaide, Australia, pp 71–81

Holland JB, Helland SJ, Sharopova N, Rhyne DC (2001) Polymorphism of PCR-based markers targeting exons, introns, promoter regions, and SSRs in maize and introns and repeat sequences in oat. Genome 44:1065–1076

Hunter DE II, Murphy JP, Phillips TD (1995) Isozyme variation in *Avena sterilis* L. collected in Turkey. Crop Sci 35:1477–1482

Jannink J-L, Gardner SW (2005) Expanding the pool of PCR-based markers for oat. Crop Sci 45:2383–2387

Karp A (2002) The new genetic era: will it help us in managing genetic diversity? In: Engels JMM, Rao VR, Brown AHD, Jackson MT (eds) Managing plant genetic diversity. International Plant Genetic Resources Institute, Rome, pp 43–56

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

Li CD, Rossnagel BG, Scoles GJ (2000) The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. Theor Appl Genet 101:1259–1268

Liu ZW, Biyashev RM, Saghai Maroof MA (1996) Development of single sequence repeat markers and their integration into a barley linkage map. Theor Appl Genet 93:869–876

Martens JW (1985) Oat stem rust. In: Roelfs AP, Bushnell WR (eds) The cereal rusts, vol II. Academic, New York, pp 103–129

Martens JW, McKenzie RIH, Harder DE (1980) Resistance to *Puccinia graminis avenae* and *P. coronata avenae* in the wild and cultivated *Avena* populations of Iran, Iraq, and Turkey. Can J Genet Cytol 22:641–649

McKenzie RIH, Harder DE (1995) Oat. In: Slinkard AE, Knott DR (eds) Harvest of gold: the history of field crop breeding in Canada. University of Saskatchewan, Saskatoon, pp 98–112

McKenzie RIH, Martens JW, Brown PD, Harder DE, Nielsen J, Boughton GR (1981) Registration of Fidler oats. Crop Sci 21:632–633

McKenzie RIH, Brown PD, Martens JW, Harder DE, Nielsen J, Gill CC, Boughton GR (1984) Registration of Dumont oats. Crop Sci 24:207

McMullen MS, Patterson FL (1992) Oat Cultivar Development in the USA and Canada. In: Marshall HG, Sorrells ME (eds) Oat science and technology. American Soc Agron Inc. and Crop Sci Soc America Publishers, Madison, pp 573–612

Murphy JP, Phillips TD (1993) Isozyme variation in cultivated oat and its progenitor species *Avena sterilis* L. Crop Sci 33:1366–1372

Pal N, Sandhu JS, Domier LL, Kolb FL (2002) Development and characterization of microsatellite and RFLP-derived PCR markers in oat. Crop Sci 42:912–918

Phillips TD, Murphy JP, Goodman MM (1993) Isozyme variation in germplasm accessions of the wild oat *Avena sterilis* L. Theor Appl Genet 86:54–64

Rajhathy T, Thomas H (1974) Cytogenetics of oats (*Avena* L.). The Genetics Society of Canada, Misc Publ No 2

Rezai A, Frey KJ (1988) Variation in relation to geographic distribution of wild oats—seed traits. Euphytica 39:113–118

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Rohlf FJ (1997) NTSYS-pc 2.1. Numerical taxonomy and multivariate analysis system. Exeter Software, Setauket

Roussel V, Koenig J, Bechert M, Balfourier F (2004) Molecular diversity in French bread wheat accessions related to temporal trends and breeding programmes. Theor Appl Genet 108:920–930

Saha MC, Rouf Mian MA, Eujayl I, Zwonitzer JC, Wang L, May GD (2004) Tall fescue EST–SSR markers with transferability across several grass species. Theor Appl Genet 109:783–791

SAS Institute Inc (2004) The SAS system for windows V8.02. SAS Institute Incorporated, Cary

Schoen DJ, Brown AHD (1993) Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. Proc Natl Acad Sci USA 90:10623–10627

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ Kansas Sci Bull 38:1409–1438

Somody CN, Nalewaja JD, Miller SD (1984) Wild oat (*Avena fatua* and *Avena sterilis*) morphological characteristics and response to herbicides. Weed Sci 32:353–359

Swofford DL (1998) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland

Takeda K, Frey KJ (1977) Growth rate inheritance and association with other traits in backcross populations of *Avena sativa* × *A. sterilis*. Euphytica 26:309–317

Thomas H (1995) Oats. In: Smartt J, Simmonds NW (eds) Evolution of crop plants. Longman Group, New York, pp 132–140

Wahl I (1958) Studies on crown rust and stem rust on oats in Israel. Bull Res Counc Isr Sec D 6:145–166

Wahl I (1970) Prevalence and geographic distribution of resistance to crown rust in *Avena sterilis*. Phytopathology 60:746–749

Wang ML, Barkley NA, Yu J-K, Dean RE, Newman ML, Sorrells ME, Pederson GA (2005) Transfer of simple sequence repeat (SSR) markers from major cereal crops to minor grass species for germplasm characterization and evaluation. Plant Genet Res 3:45–57

Wu X, Larson SR, Hu Z, Palazzo AJ, Jones TA, Wang RR-C, Jensen KB, Chatterton NJ (2003) Molecular genetic linkage maps for allotetraploid *Leymus* wildryes (Gramineae: Triticeae). Genome 46:627–646

Yu J-K, Dake TM, Singh S, Benscher D, Li WL, Gill B, Sorrells ME (2004) Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. Genome 47:805–818

Zeven AC, Zhukovsky PM (1975) Dictionary of cultivated plants and their centres of diversity. Centre for Agricultural Publishing and Documentation, Wageningen

Zhou X, Jellen EN, Murphy JP (1999) Progenitor germplasm of domesticated hexaploid oat. Crop Sci 39:1208–1214

Zohary D, Hopf M (2000) Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley, 3rd edn. Oxford University Press, Oxford